LE RÔLE DES STATISTIQUES DANS L'APPRENTISSAGE DU LANGAGE.

JACQUES MEHLER

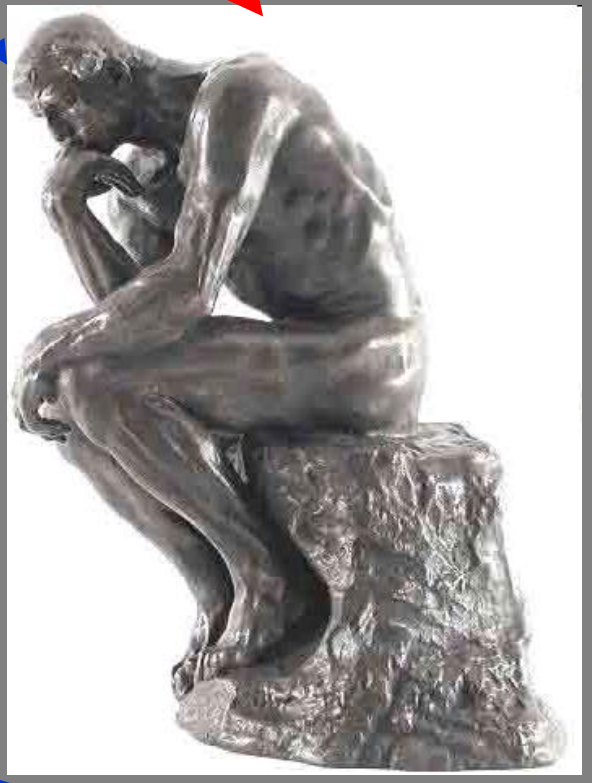
Paris, Février 16, 2010

- Theories and facts about early language acquisition.
- Core properties with statistics vs stats without core properties?
- First, a little detour.

Most of our investigations in the last 10 years have focused on neonate cognition and very young infants.

Why should one study language dispositions in the newborns?

These two animals are quite similar but only this one speaks …



Both animals have the ability to extract statistical information.

Why should one be interested in the study of language dispositions in the newborns?

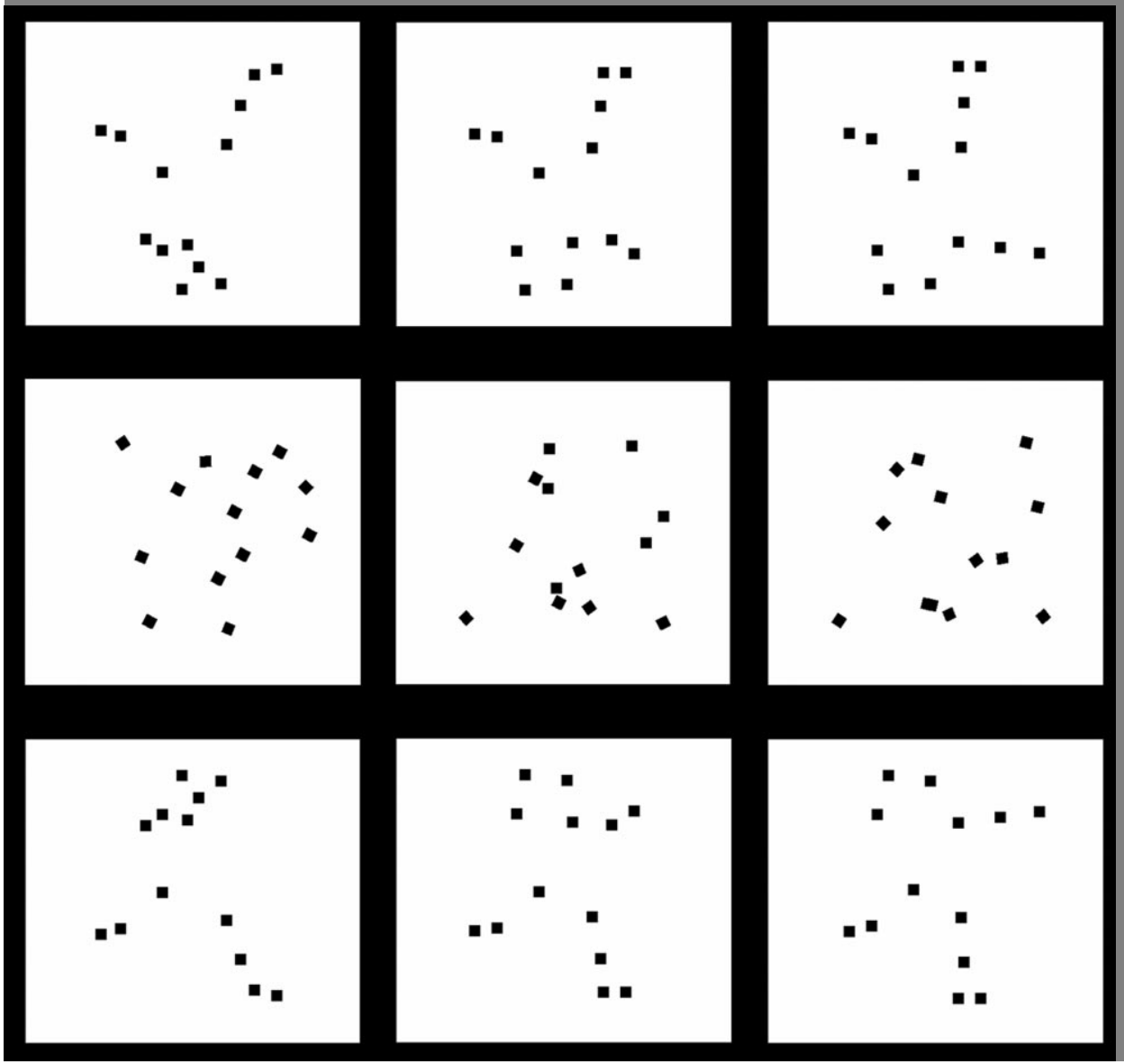
Older infants are easy to test. Neonates are not easy to study. But some of us decided to go ahead anyhow.

The discoveries of Ethologists as Lorenz, Tinbergen, Marler, & Nottebohm made it clear that, as Darwin claimed, ancestry is fundamental for understanding the behaviors of different species.

Lorenz discovered that ducks and geese *IMPRINT*

… do all animals imprint? Probably not, but ducks imprint to the first moving object that it sees.

BIOLOGICAL MOTION IS AN IMPORTANT CONCEPT.



Cutting et al., Vallortigara et al. demonstrations showing that biological motion is essential to understand how not only chicks and other animals are attracted to biological motion. Also neonates are sensitive to B.M.

Would ducks imprint to a car, or another man made artifact?

F.Simion et al. PNAS (2009)

What would have happened if K. Lorenz had studied 3-months old ducks?

He would have failed...to discover imprinting!

- Ancestry determines specialized behaviors of animals.

- Are humans animals born with a single specificity: the ability to learn?

- If we believe in learning only, it is obvious why psychologists had neglected language acquisition, music cognition, etc. (until the sixties).

- Neonates have not had time to acquire knowledge…but they may have core linguistic, musical and other cognitive dispositions, justifying the study of the neonates.

W. James convinced students of development that the mind of newborns is in a state of chaos. His argument was that the regularities present in the environment gradually “shape” the development of the mind.

BUT…

Developmental studies in Paris and several other locations promoted a change of “weltanschauung”.

The work that is relevant to our quest will begin with the studies with older infants and even adults

- In 2001 we organized a conference at SISSA on cognitive development. Two trends appeared as major players in the field. One, held the view that distributional evidence was sufficient for infants to learn their native language and other cognitive skills. The other claimed that a rule extraction mechanism is essential to learn the syntax of the native language.

Saffran et al. (1996) had discovered that 8-month- olds segment a speech stream.

- The meaningless, prosodic free, stream contained *tri-syllabic statistical* words. That is: the 1st syllable of a word predicts the 2nd, and it predicates the 3rd. Third syllables are poor predictors of the next word

• Saffran et al.’s work gave rise to the view that an all statistical account of linguistics can do away with Generative Grammar accounts as Chomsky and his colleagues had proposed, see Elman & Bates, Science (1996).

• Marcus et al (1999) in a study with 7-month-olds showed that they could extract regularities carried by discrete items. He used three kinds of “grammars”: AAB, ABB and ABA.

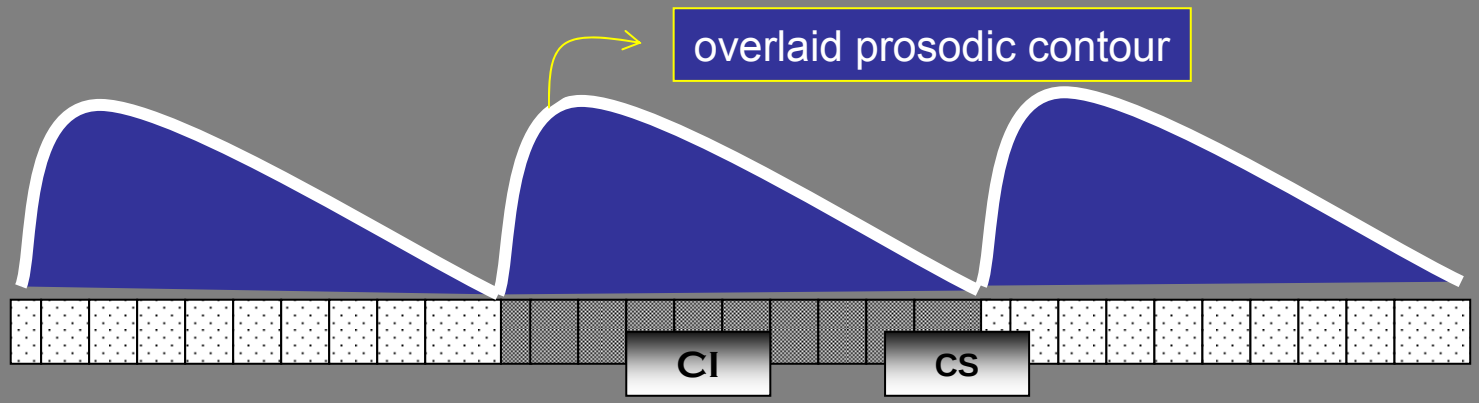

• Both studies suggested that infants learn language at a pre-lexical stage.

• However, I will argue that these elegant studies might be unsuitable to accept their claims.

• Consider first the work of Saffran et al. They relied on streams that are non-speech like. If learning to speak is an instinct as (see Garrett) then such stimuli will not reveal what infants do in the household situation.

In his thesis M. Shukla explored language acquisition introducing prosodic structures.

Contrasting Contour-Internal and Contour-Straddling 'words'.



CI : Contour-internal word
CS : Contour-straddling word

Shukla, Nespor & Mehler (2007)

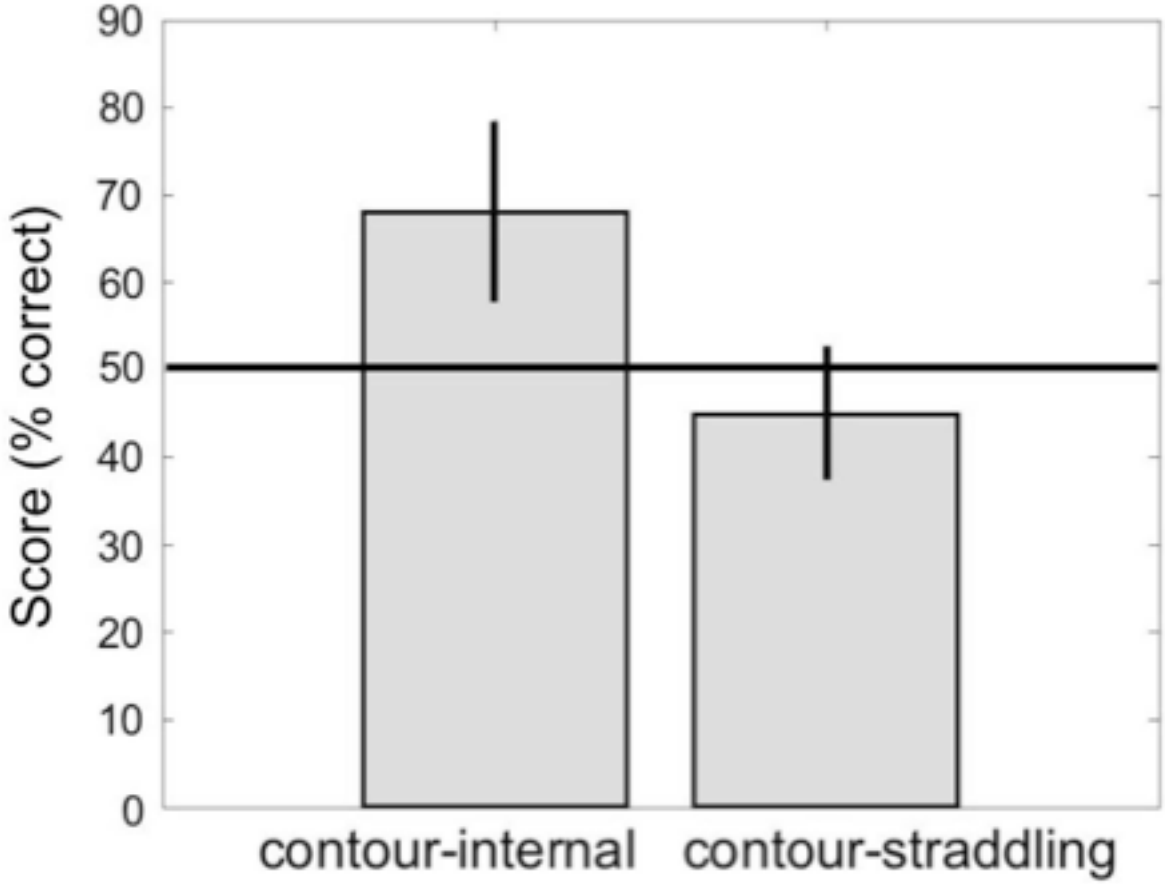
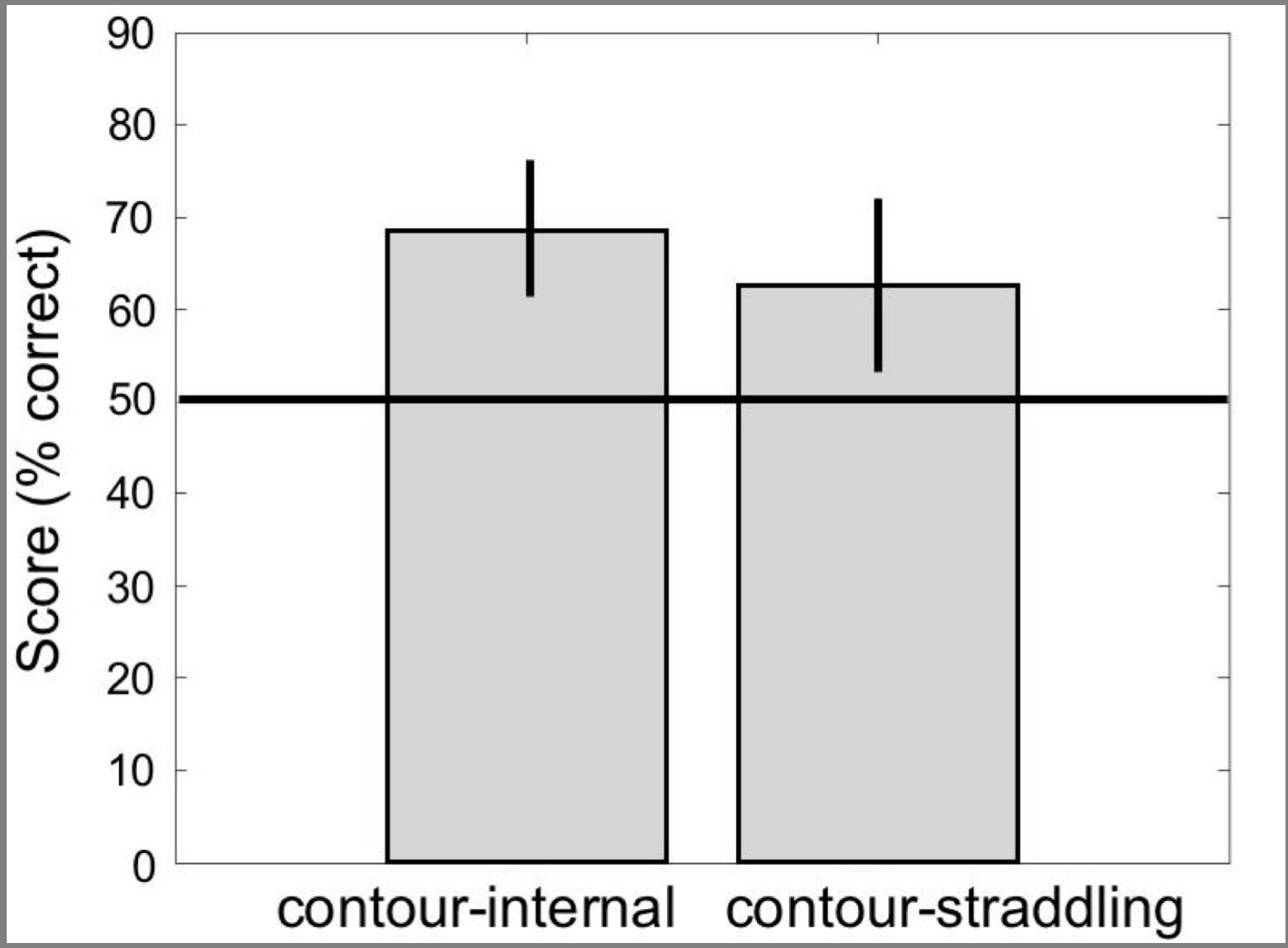


Figure 3. The mean scores (% correct) from experiment 2. Only contour-internal 'words' appear to be correctly segmented. Error bars represent 95% confidence limits of the means.

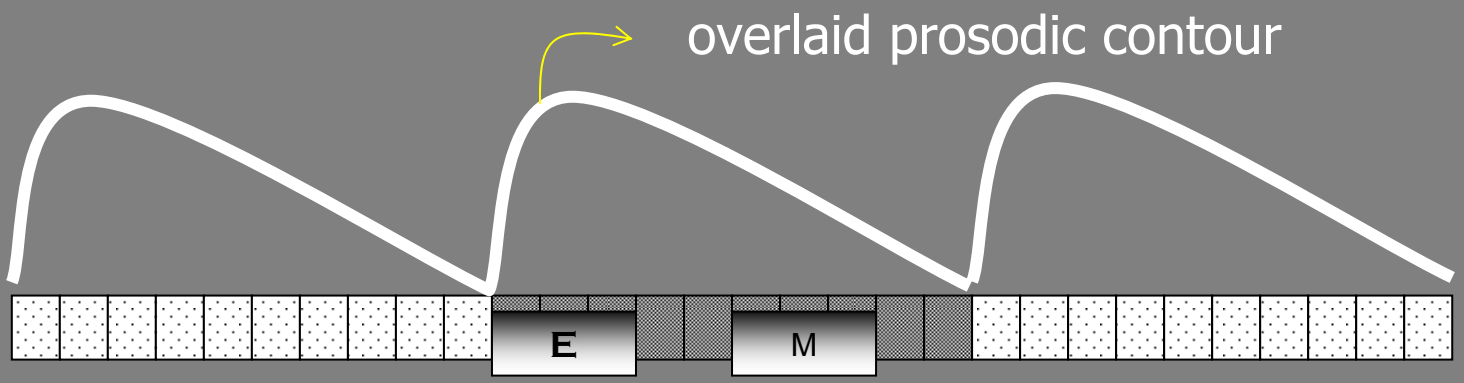
CONTROL WITHOUT PROSODIC COUNTOURS



• Endress, Scholl & Mehler (2005) demonstrated that edges are highlighted positions that permit further computations.

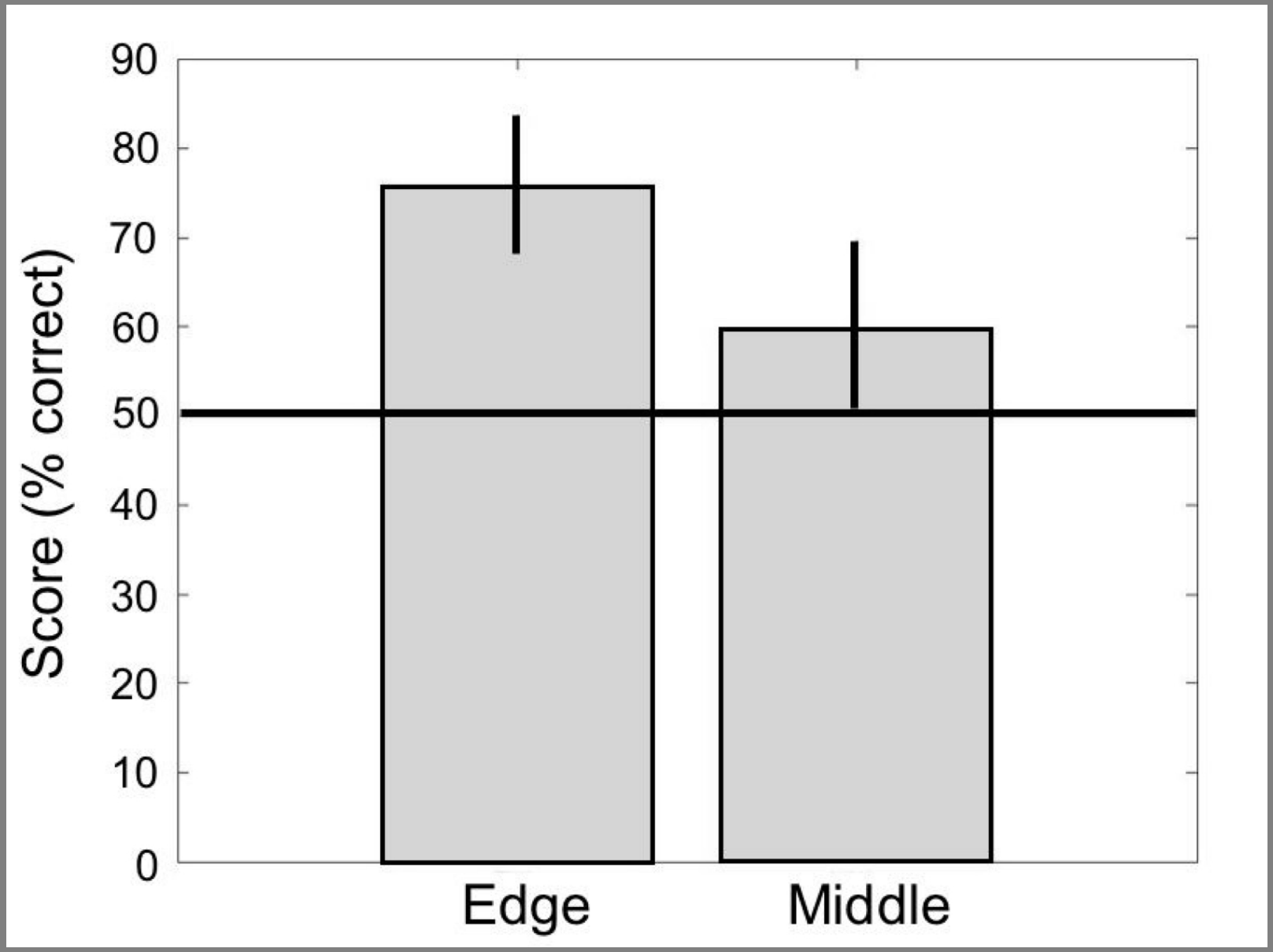
•Shukla et al. made use of this finding to validate the role of prosody.

Contrasting Contour-Edge and Contour-Middle 'words'.

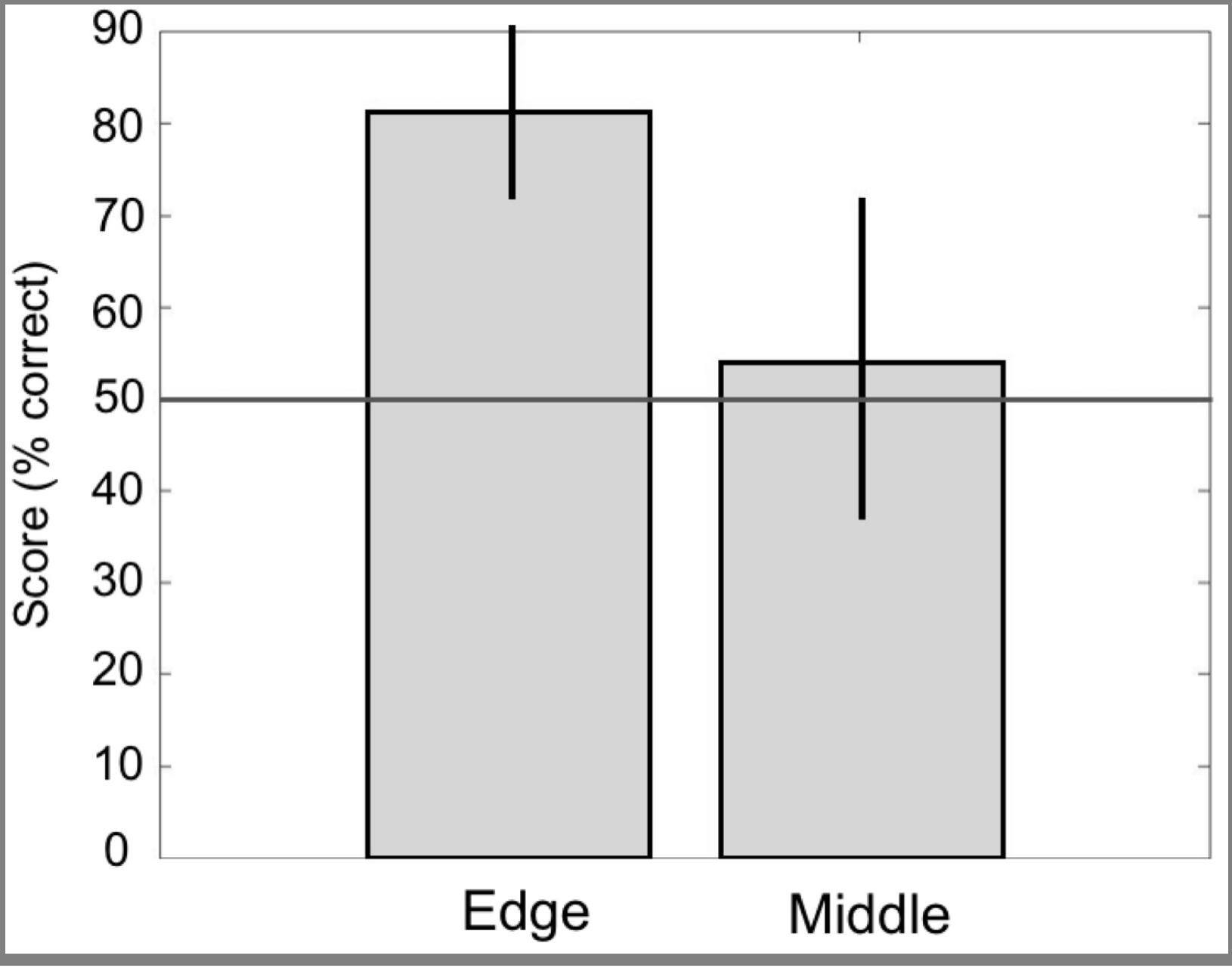


E : 'Edge' word
M : 'Middle' word

LEFT EDGE VERSUS MIDDLE



RIGHT EDGE VERSUS MIDDLE



Shukla’s work suggested that whenever possible one should use controlled but realistic speech stimuli, rather than the monotone speech streams that had become so popular with most psycholinguists.

Despite the limitation, stats computations play an important role in language acquisition. So do rules. But I believe that Marcus' demo has also problems. There is no time to review these, however, we go through these problems in a recent paper:
Endress, Nespor & Mehler (2009) *Trends In Cognitive Science*. Perceptual and memory constraints on language acquisition.

All the pdfs of the papers we published can be downloaded from

www.sissa.it/cns/lcd/publications.html

Both, stats and rules play an important role in language acquisition. How do they interact ?

Pena et al. (2002) explored this issue:

- Exp Saffran et al. replicated using AxC non-adjacent TPs. Adults segment the stream into its “words”.
- Exp 2: Ss do not generalize a “if A then C” rule, no pauses.
- Exp 3: Ss generalize with 25 ms pauses after each C.
- Exp 4: No generalization with 30 min familiarization, no pauses.
- Exp 5: Generalization when familiarization reduced to 2 min, and pauses.

These experiments show that the two mechanisms, stat extraction and generalization operate under conditions that are determined by properties of the signals.

The notion that there are constraints that determine the when and where statistical mechanisms operate does not mean that they do not play an important role in language acquisition

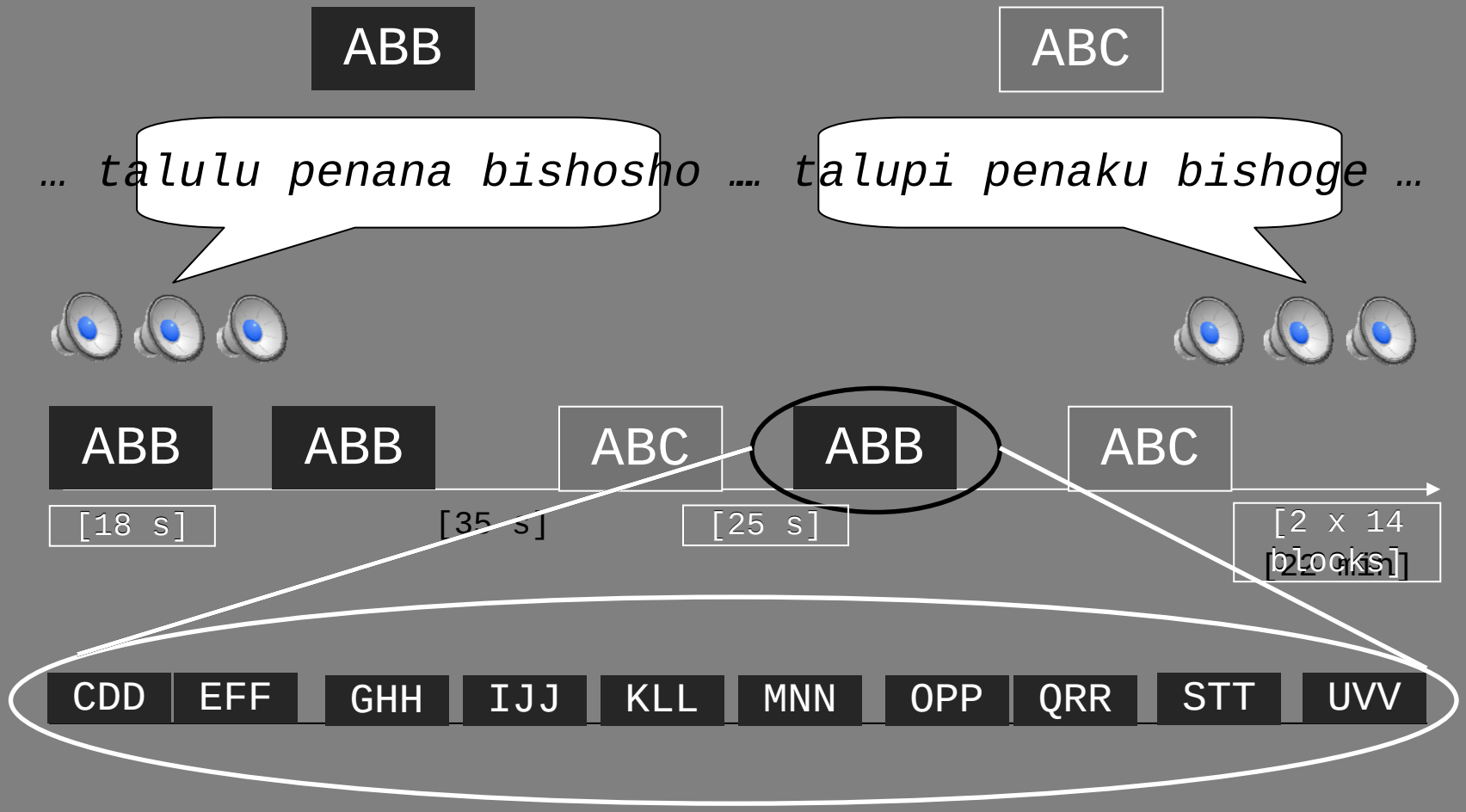
Likewise, rules intervene in the acquisition of grammar. However, the demonstration of Marcus is not entirely convincing. In fact, the problem is that two of the constraints that we were able to establish is the important role of repetitions and of edges.

Gervain in her thesis explored whether the neonate is capable of behaving as Marcus et al'. 7-month-olds.

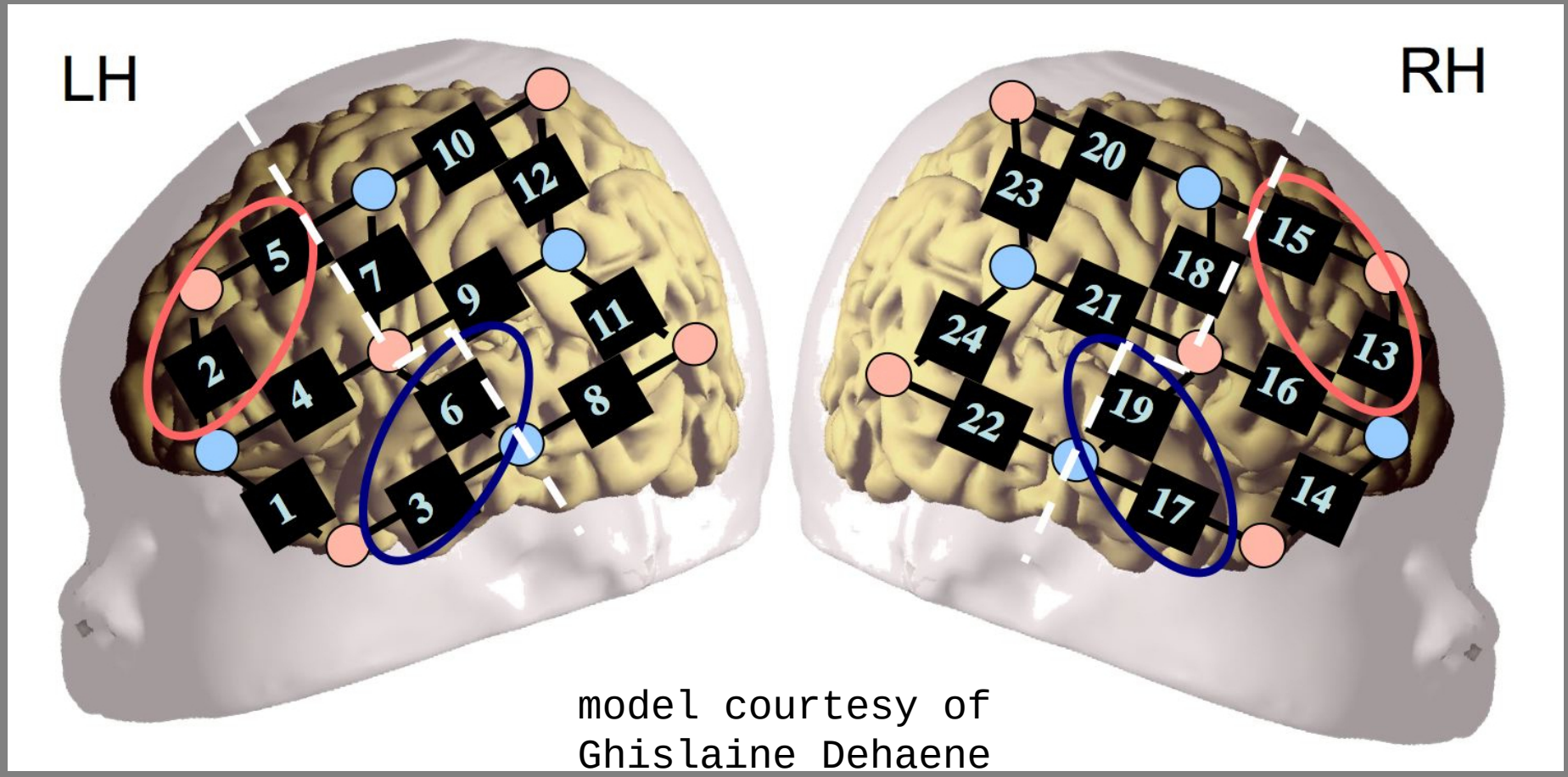
In a NIRS based study with infants, Gervain et al. (2008) tested neonates, using a 24-channel device that records hemodynamic responses from 12 areas of the L-hemisphere and 12 areas from the R-hemisphere.

A neonate being prepared to participate in a Near Infra-Red Spectroscopy (NIRS) experiment .

Can infants detect syllable duplication



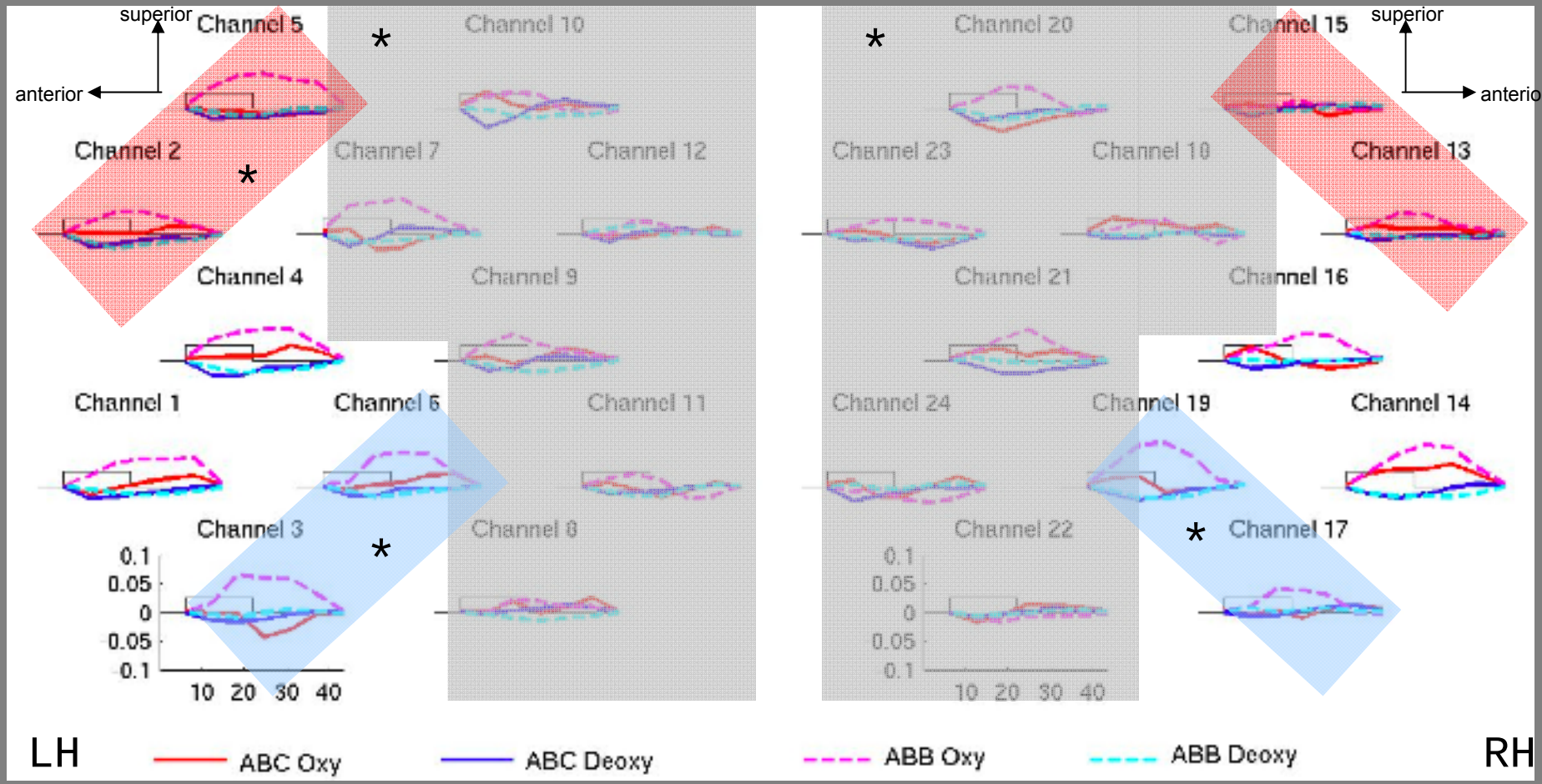
Method



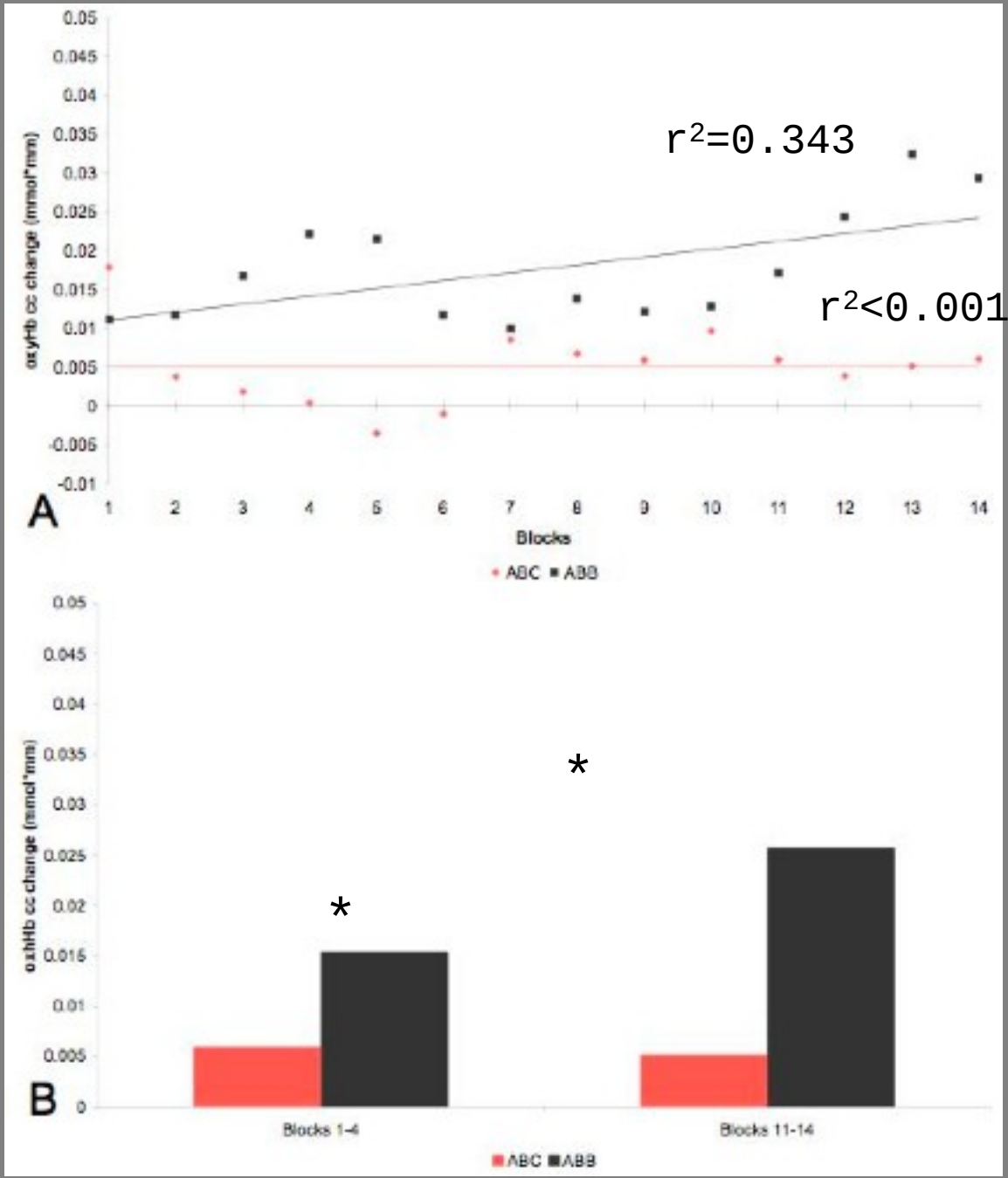
model courtesy of
Ghislaine Dehaene

Results

n = 22


superior
anterior
Channel 5
Channel 10
Channel 2
Channel 7
Channel 12
Channel 4
Channel 9
Channel 1
Channel 6
Channel 11
Channel 3
Channel 8
0.1
0.05
0
0.05
0.1
10 20 30 40
Channel 20
Channel 15
Channel 23
Channel 18
Channel 13
Channel 21
Channel 16
Channel 24
Channel 19
Channel 14
Channel 22
Channel 17
LH
ABC Oxy
ABC Deoxy
ABB Oxy
ABB Deoxy
RH

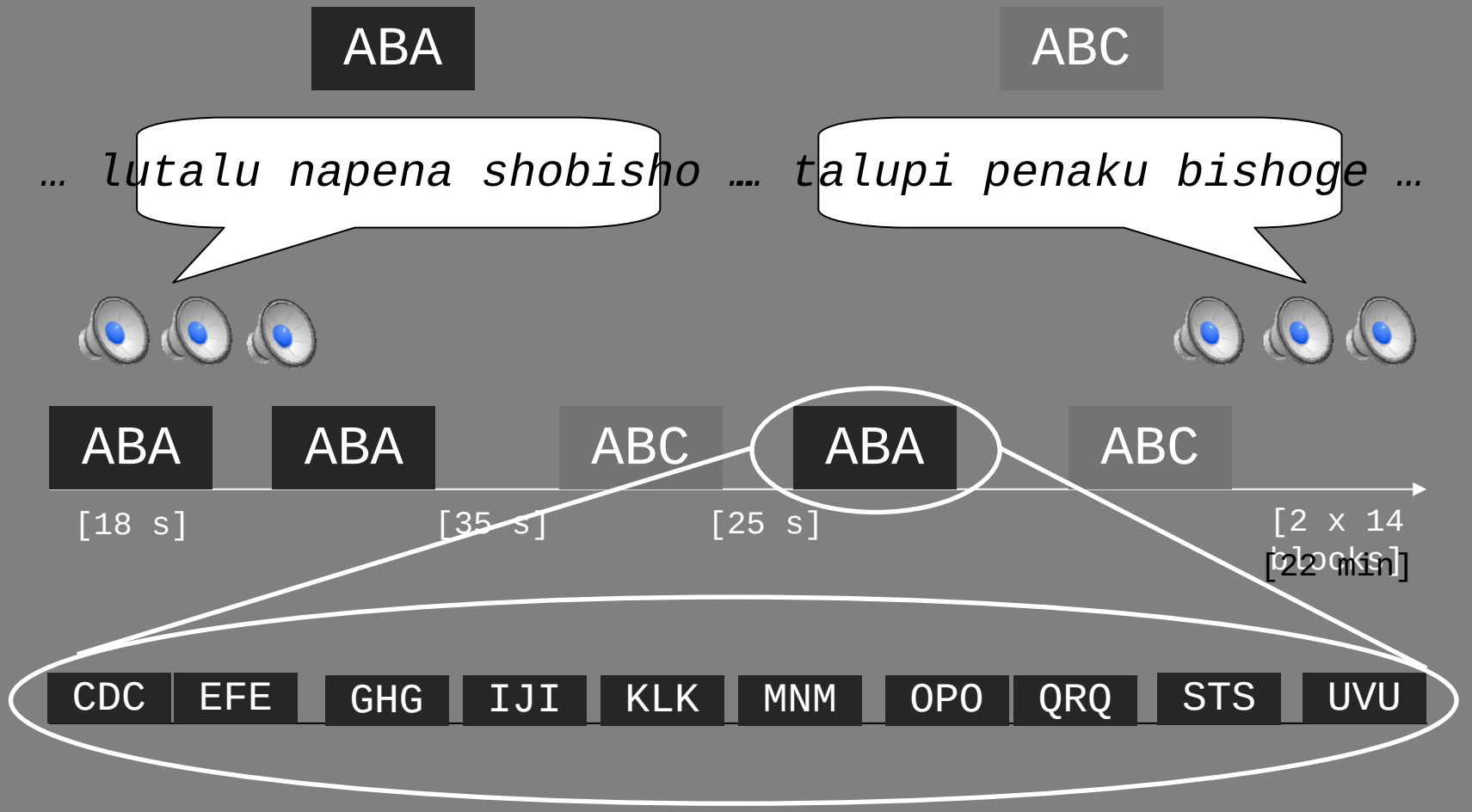
Results


r²=0.343
r²<0.001
oxyHb cc change (mmol*mm)
Blocks
ABC ABB
A
oxHb cc change (mmol*mm)
*
*
Blocks 1-4
Blocks 11-14
ABC ABB
B

EXP 2

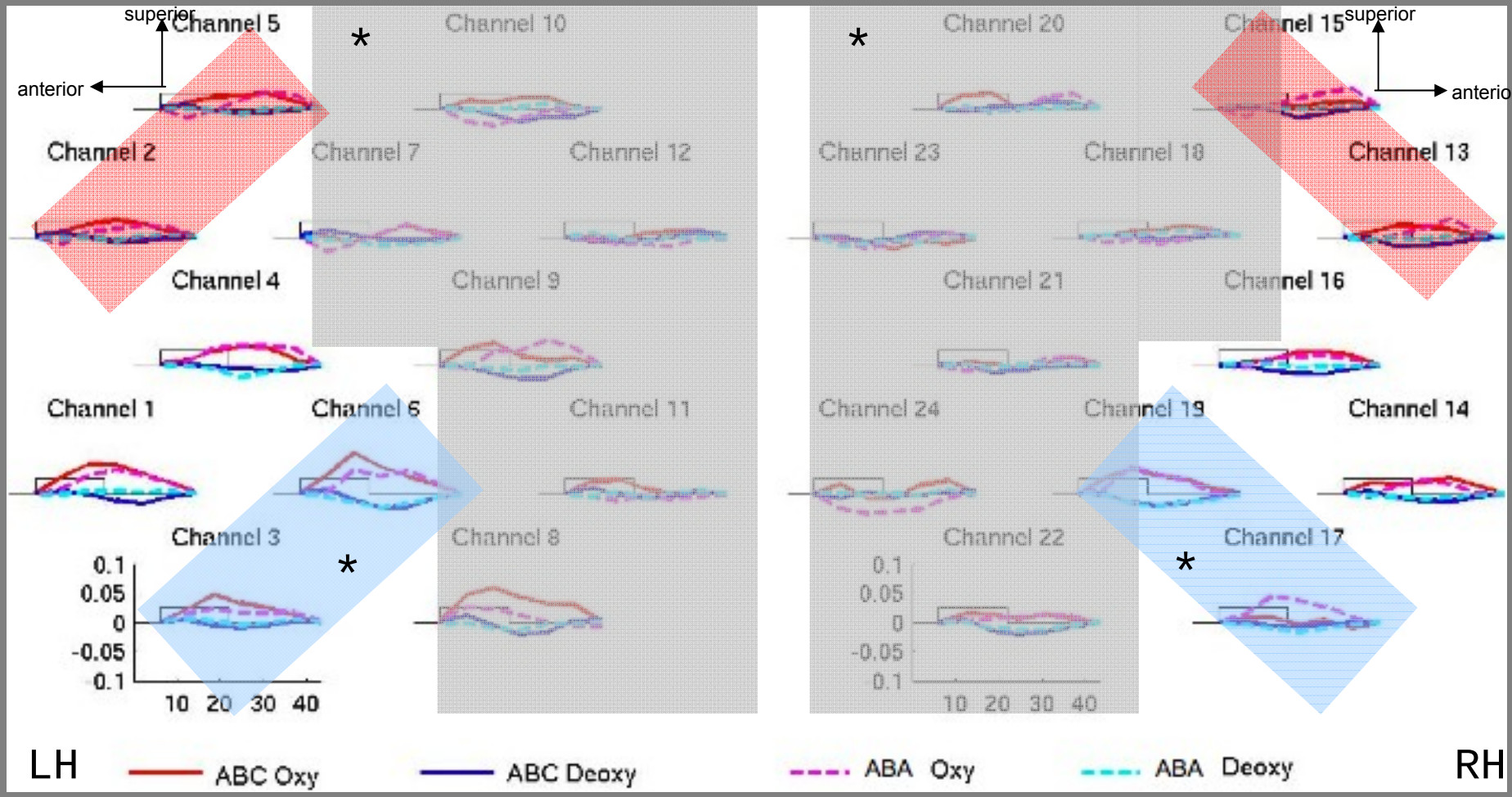
- can newborns also detect and learn non-adjacent dependencies?

Non-adjacent repetition test

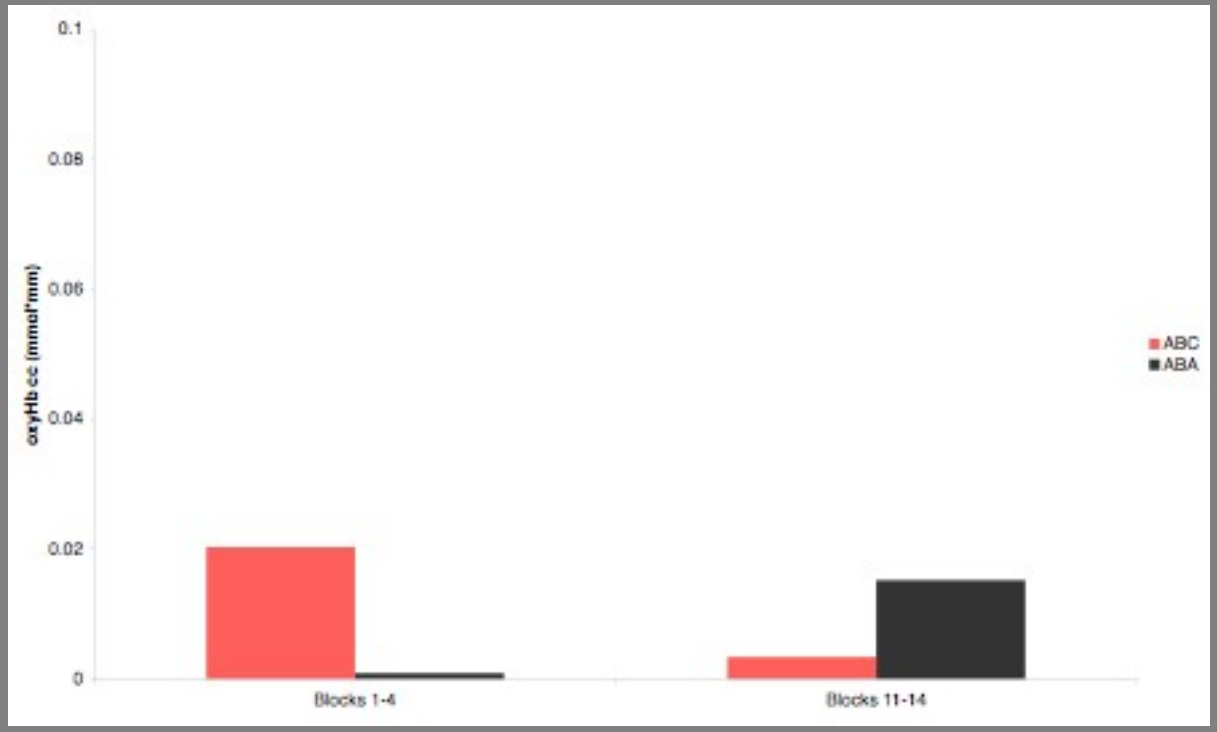


Results

n = 22


superior
anterior
Channel 5
Channel 10
Channel 2
Channel 7
Channel 12
Channel 4
Channel 9
Channel 1
Channel 6
Channel 11
Channel 3
Channel 8
0.1
0.05
0
-0.05
0.1
10 20 30 40
Channel 20
Channel 15
Channel 23
Channel 18
Channel 13
Channel 21
Channel 16
Channel 24
Channel 19
Channel 14
Channel 22
Channel 17
LH
ABC Oxy
ABC Deoxy
ABA Oxy
ABA Deoxy
RH

Results


0.1
0.08
0.06
0.04
0.02
0
oxyHb cc (mmol*mm)
Blocks 1-4
Blocks 11-14
ABC
ABA

Summary

- newborns detect adjacent repetitions as perceptual Gestalts, similarly to adults.
- non-adjacent repetitions, however, are not detected.
- newborns learn regularities based on these perceptually salient patterns.
- the neonate brain shows functional specification. (Peña et al 03, Dehaene-Lambertz et al 02, 06)

Let me now turn to other notions that play important roles in development.

Two decades ago we discovered that infants compute relations many months before they remember words (or objects).

In studies of neonate sensitivity of linguistic rhythm, Ramus, Nespor and others, postulated that infants handle vowels and consonants as different objects. We should have talked about *categories*.

The two categories have been recognized since at least 2.500 years by careful observers, I have in mind Vowels and Consonants. Pāṇini, (4 th BC century) a Sanskrit linguist and one of the earliest descriptive and generative grammarians, distinguished vowels from consonants.

Spinoza nearly 350 years ago wrote that:

"Ce sont les accents, leur mélodie entraine en mouvement derriere eux, comme une armée derriere le Roi, les lettres et les voyelles. Les lettres, ce sont le corps; les voyelles, sont l'âme; toutes suivent la marche des accents et s'arretent en meme temps qu'eux"[3]. We quote from the French version entitled *Abrégé de Grammaire Hébraïque*, published in Paris in 1968. Spinoza's (1677) insights are being vindicated by numerous recent empirical reports, as we shall see below.

Mehler, Bonatti, Peña & Nespor, *Cortex, 2006*

Nespor, Peña & Mehler (2003) in a theoretical paper argued that these two categories play different roles. The Cs are mostly central for lexical purposes, whereas the Vs are relevant to detect grammatical regularities.

SEVERAL LANGUAGES HAVE C-ROOTS

Gdl :"big" not a word, just a root

gadol :"big" masculine adjective
gdola : "big" (feminine adjective)
giddel : "he grew" (transitive verb)
gadal : "he grew" (intransitive verb)
higdil : "he magnified" (transitive verb)
magdelet: "magnifier" (lens)

Sfr : root of "count" or "recount"

Sefer : "book" (containing tales which are recounted)
sofer : "scribe" (Masoretic scribes; counted verses)

WE DON'T KNOW OF LANGUAGES THAT HAVE VOWEL ROOTS

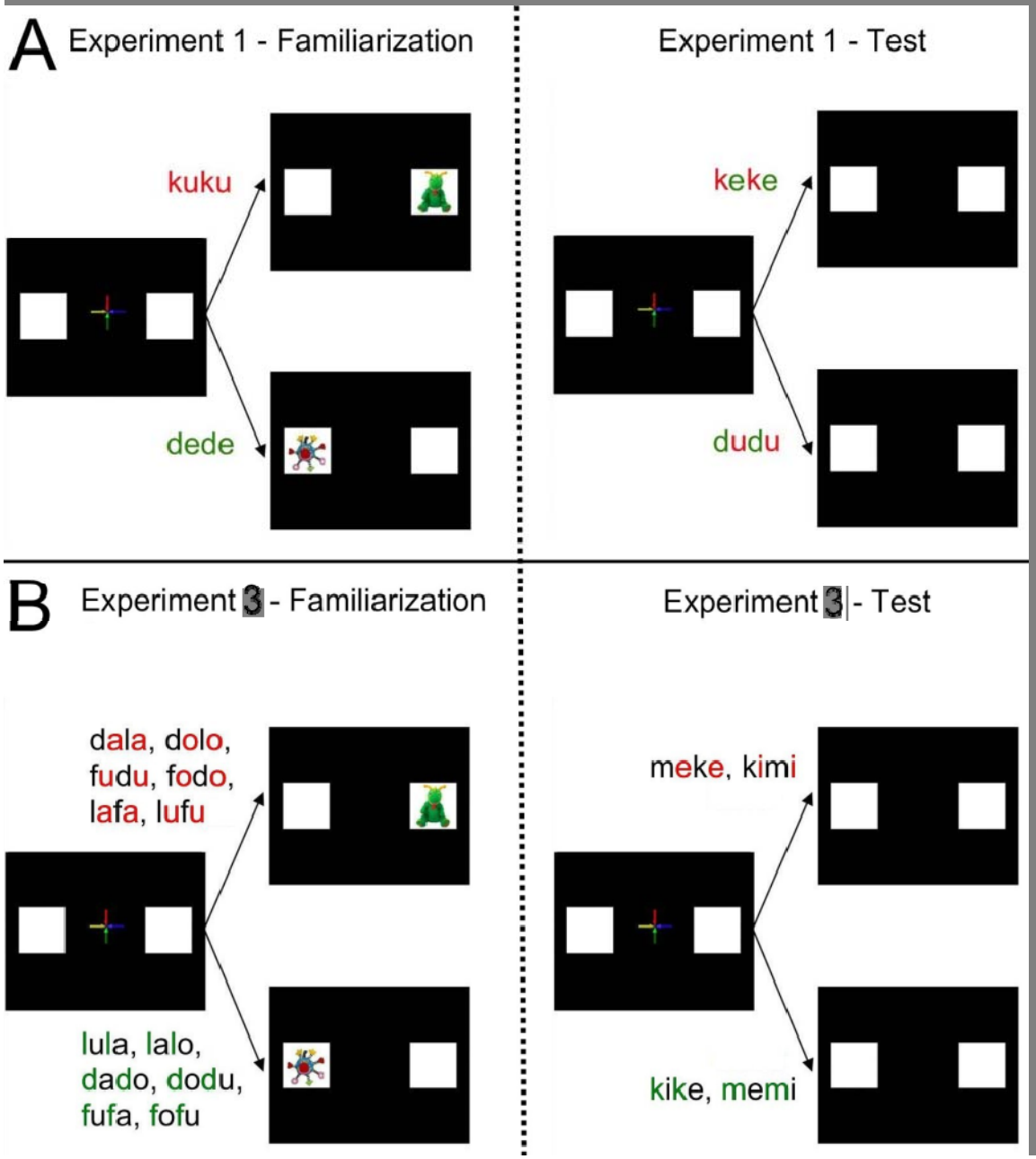
A number of empirical studies followed Nespor et al. (2003):
Bonatti, Peña, Nespor & Mehler
Psych. Sci. (2005)
Mehler, Peña, Nespor & Bonatti Cortex (2006),

The above papers established that:
C-sequences are best to individuate lexical items and are used to compute stats.
V-sequences are best to extract regularities and no stats are computed over them.

Keidel et al. (2005) argued that our results could be a consequence of the fact that in all languages more Cs are used that Vs. One could ask “why does this regularity arise universally” ? In order to compute such statistics it would be necessary to have at least some memory allowing to compute that in the particular native language this regularity exists.

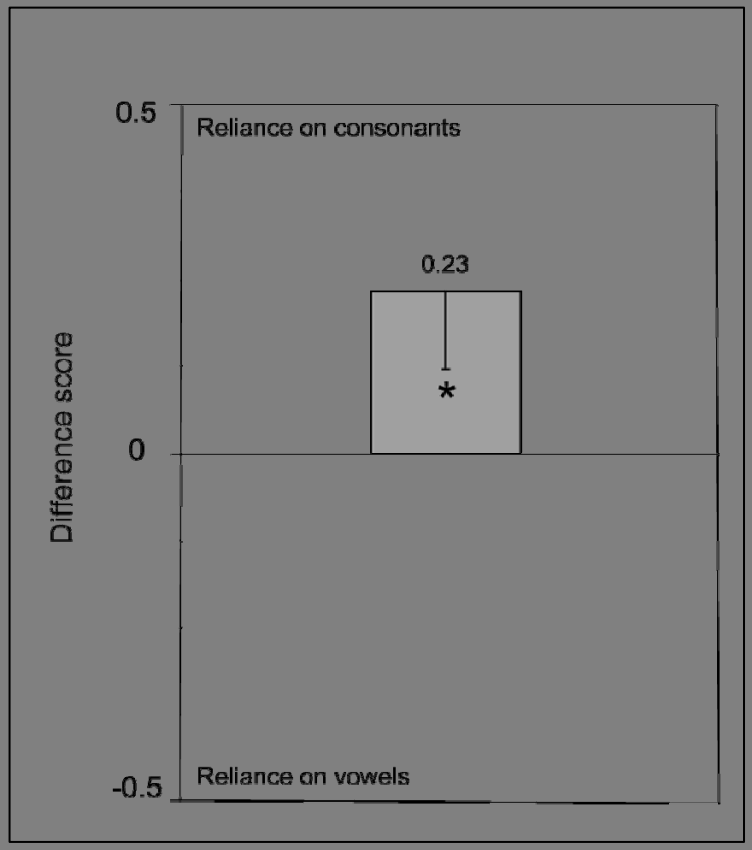
To address Keidels conjecture, Hochmann, Benavides, M. Nespor & Mehler (submitted) explored whether very young infants behave like adults.

Procedures in Hochmann, Benavide, Nespor & Mehler (submitted): 12 months olds. N26 in each experiment.

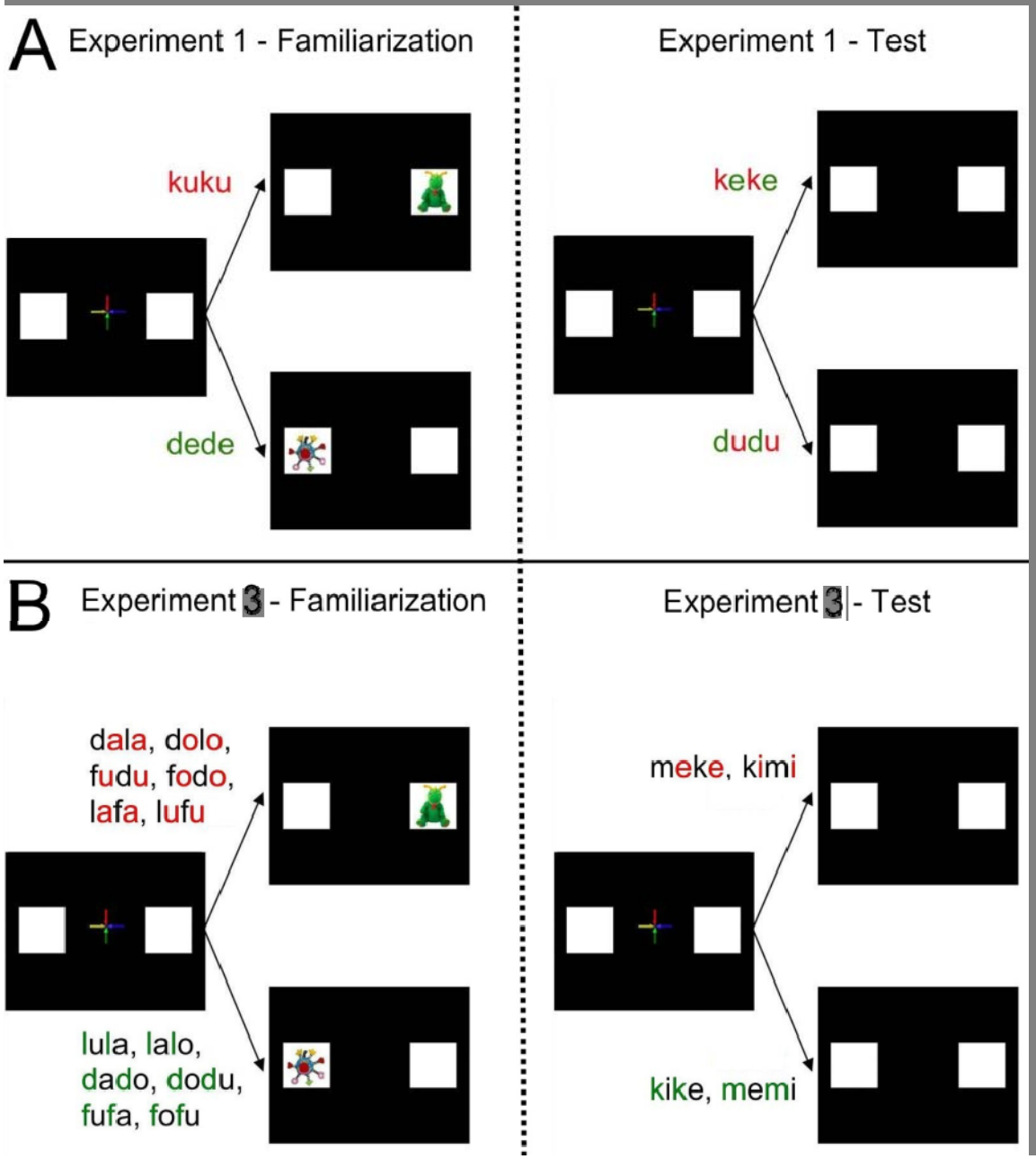


Difference scores:
(#consonant looks - #vowel looks) / (#consonant looks + #vowel looks)

The results were significant on a t-test and moreover, 17 had positive scores, 6 had negative scores and six had null-scores.

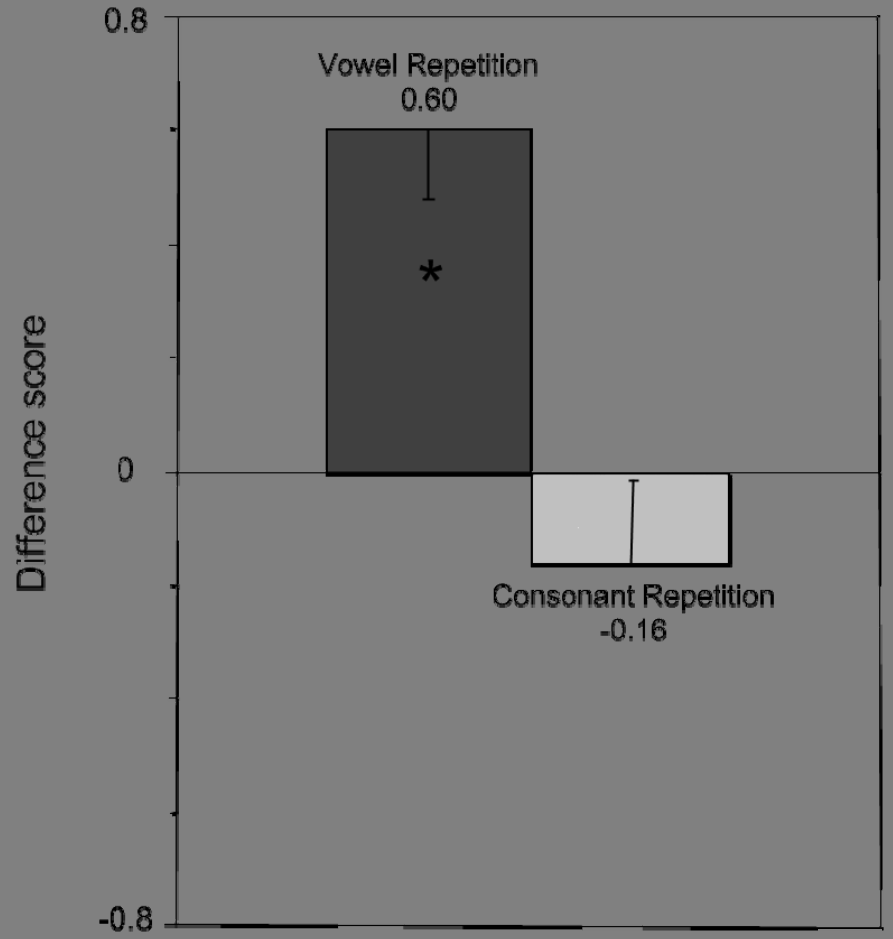


Procedures in Hochmann, Benavide, Nespor & Mehler (submitted): 12 months olds. N26 in each experiment.



Difference scores:
(#consonant looks - #vowel looks) / (#consonant looks + #vowel looks)

In experiment 2, N=24, participants: 12-month-olds
Results are highly significant: For vowel repetition P<.0001;
d’ = 2.147 . No significant results for C-repetition



These results were observed at an age when infants have 80 or more lexical items.

In experiments that are in progress with pre-lexical participants, the results show a very similar trend.

Let me finish with some studies in progress that allow me to satisfy a curiosity of mine. When I first arrived in Paris I engaged in a study of memory development.

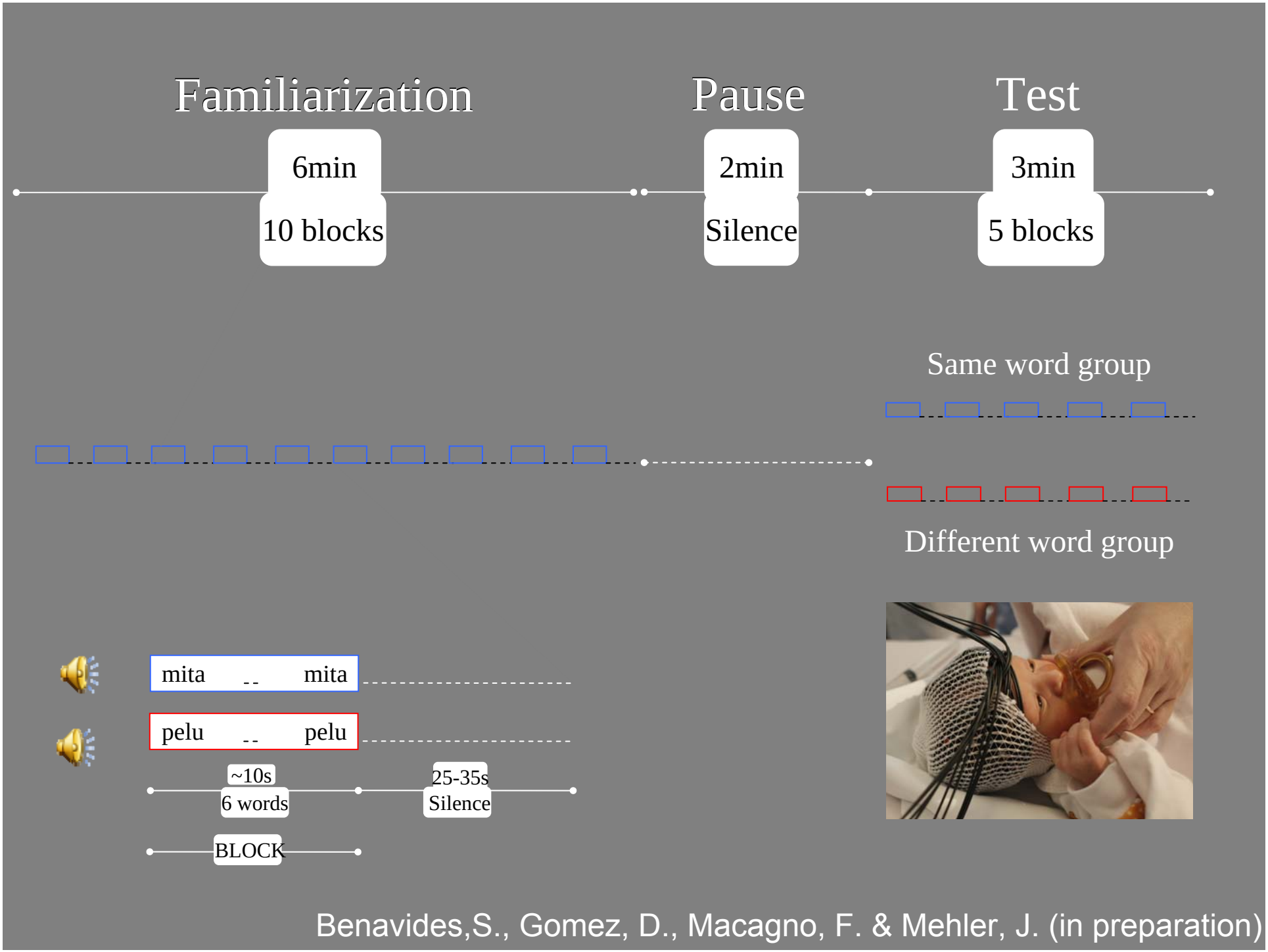
I am now revisiting this issue for two reasons:

First: Why is it that neonates do not display memory for frequently heard words?

Second: Are the representations of language and music encoded alike?

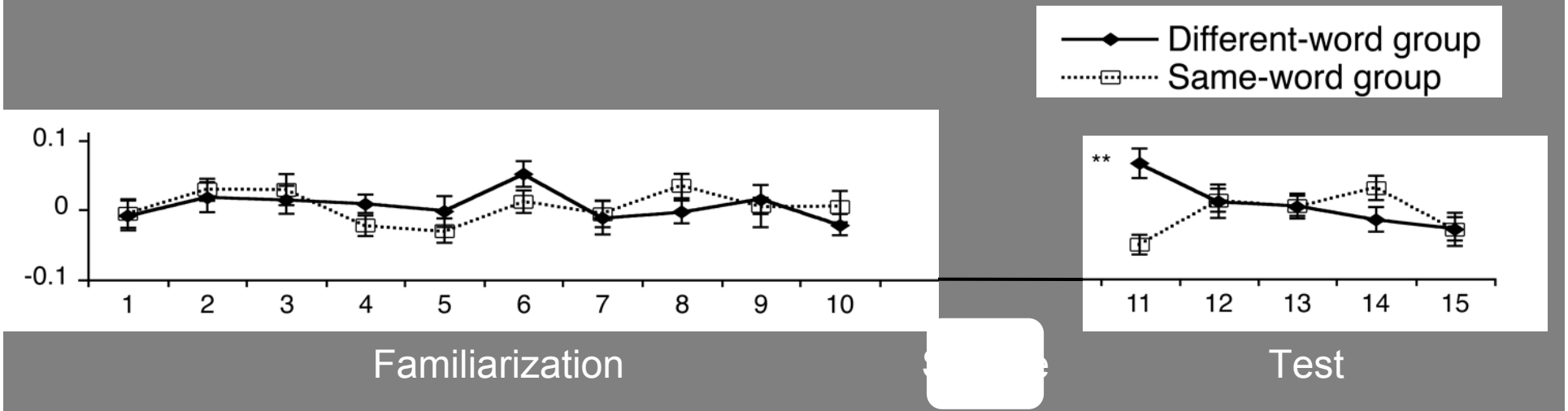
Benavides, Gómez et al. are now testing
which properties of stimuli are encoded

Are identity relations established:
over features, prosodic properties,
speaker identity, and/or number of syllables?

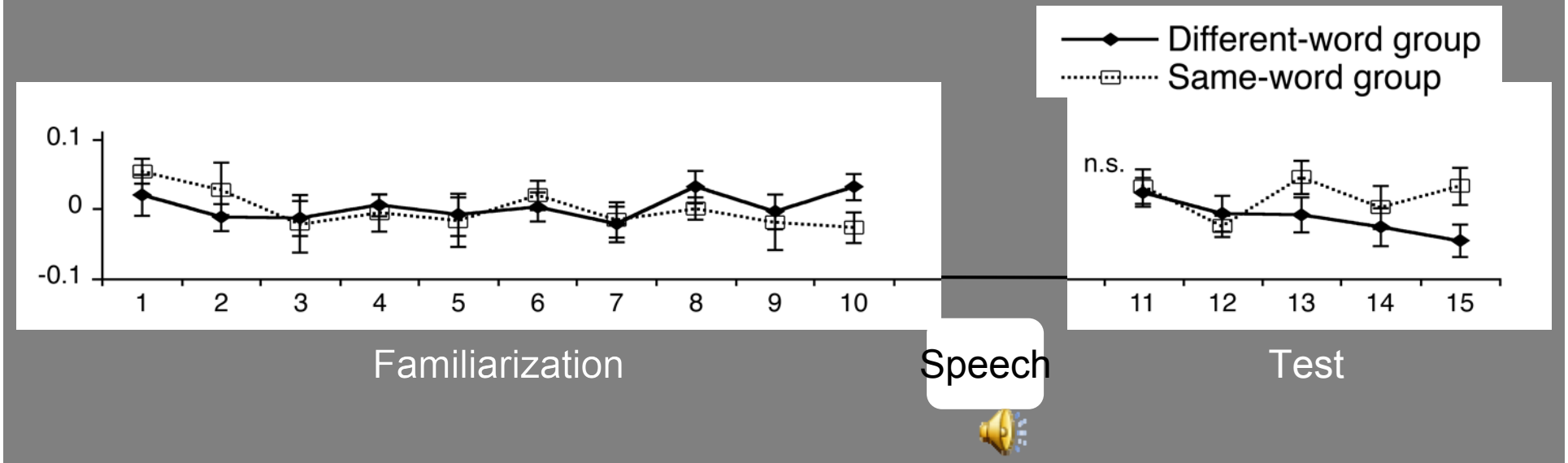


Benavides,S., Gomez, D., Macagno, F. & Mehler, J. (in preparation)

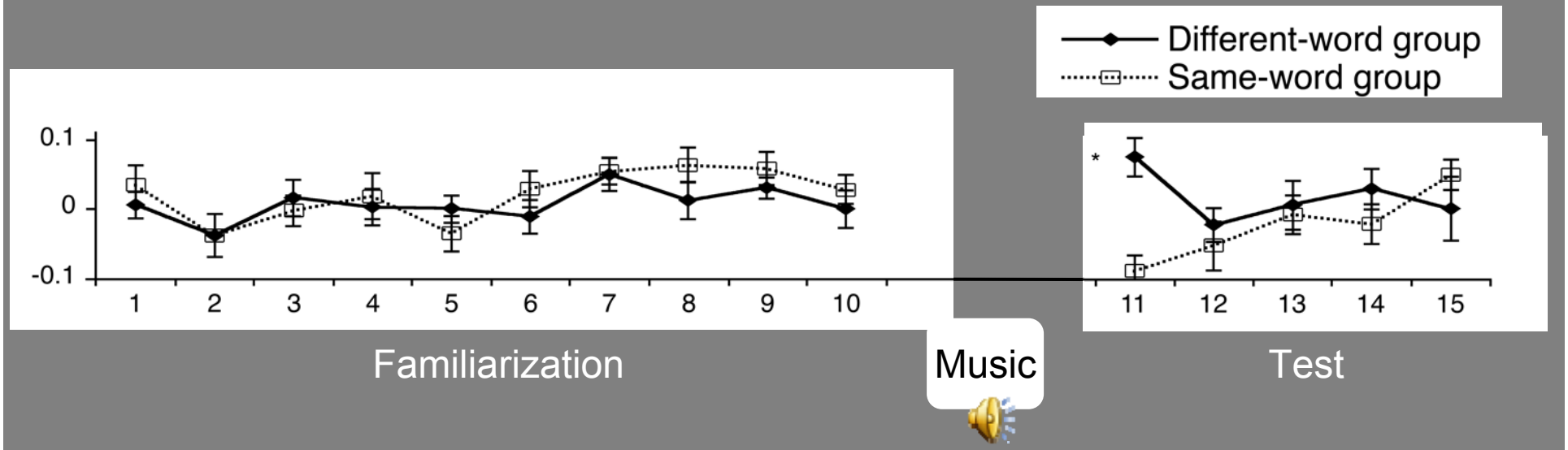
Retention of a word over a silent interval (56 neonates)



Interference of a word instead of silence (28 neonates)



Interference of a tune (Brahms lullaby) instead of silence (28 neonates)



Conclusions

- I tried to justify why, when one wants to understand language acquisition, it becomes necessary to start with very young infants.
- My purpose was to illustrate that it is also necessary to study changes that are taking place during growth, combing behavioral and imaging techniques.

Language acquisition is constrained by perceptual, memory and other cognitive functions. It is a mistake to ignore these constraints while postulating a single mechanism that does it all.

A question for future studies is whether speech and grammar constrain one another.